BOFIT Discussion Papers
18 • 2011

Peter Sarlin, Tuomas A. Peltonen

Mapping the State of Financial Stability

EUROJÄRJESTELMÄ
EUROSYSTEMET

Bank of Finland, BOFIT
Institute for Economies in Transition

BOFIT Discussion Papers
Editor-in-Chief Laura Solanko

BOFIT Discussion Papers 18/2011
27.7.2011

Mapping the State of Financial Stability

ISBN 978-952- 462-713-9
ISSN 1456-5889
(online)

This paper can be downloaded without charge from
http://www.bof.fi/bofit

Suomen Pankki
Helsinki 2011

Contents

All opinions expressed are those of the author and do not necessarily reflect the views of the Bank of Finland.

Peter Sarlin[1], Tuomas A. Peltonen

Mapping the State of Financial Stability[2]

Abstract

The paper uses the Self-Organizing Map for mapping the state of financial stability and visualizing the sources of systemic risks on a two-dimensional plane as well as for predicting systemic financial crises. The Self-Organizing Financial Stability Map (SOFSM) enables a two-dimensional representation of a multidimensional financial stability space and thus allows disentangling the individual sources impacting on systemic risks. The SOFSM can be used to monitor macro-financial vulnerabilities by locating a country in the financial stability cycle: being it either in the pre-crisis, crisis, post-crisis or tranquil state. In addition, the SOFSM performs better than or equally well as a logit model in classifying in-sample data and predicting out-of-sample the global financial crisis that started in 2007. Model robustness is tested by varying the thresholds of the models, the policymaker's preferences, and the forecasting horizon.

JEL Codes: E44, E58, F01, F37, G01.
Keywords: systemic financial crisis, systemic risk, Self-organizing maps, visualisation, prediction, macroprudential supervision

[1] Corresponding author: Peter Sarlin, Department of Information Technologies, Åbo Akademi University, Turku Centre for Computer Science, Joukahaisenkatu 3–5, 20520 Turku, Finland. Email: psarlin@abo.fi tel. +358 2 215 4670.
[2] The authors want to thank Barbro Back, Tomas Eklund, Kristian Koerselman, Marco Lo Duca and Fredrik Lucander, and seminar participants at the ESCB Macro-Prudential Research Network (MaRs) workshop on 14–15 April 2011 in Frankfurt am Main, the International Joint Conference on Artificial Intelligence (IJCAI'11) workshop on Chance Discovery on 16–22 July 2011 in Barcelona, the Bank of Finland Institute for Economies in Transition (BOFIT) and the Data Mining and Knowledge Management Laboratory at Åbo Akademi University for useful comments and discussions. All remaining errors are of our own. The views presented in the paper are those of the authors and do not necessarily represent the views of the European Central Bank or the Eurosystem.

Peter Sarlin, Tuomas A. Peltonen

Mapping the State of Financial Stability

Tiivistelmä

Tutkimuksessa hyödynnetään Itseorganisoituvaan karttaan (engl. Self-Organizing Map, SOM) perustuvaa mallia kuvaamaan rahoitusmarkkinoiden vakautta ja systeemiriskin lähteitä sekä ennustamaan finanssikriisejä. SOM-mallin avulla moniuloitteinen aineisto voidaan kuvata kaksiuloitteisesti, mikä mahdollistaa systeemiriskiin vaikuttavien yksittäisten muuttujien analysoinnin. Lisäksi mallia voidaan käyttää osoittamaan yksittäisen maan sijainti finanssivakaussyklin eri vaiheissa. Syklin vaiheita ovat kriisiä edeltävä, kriisi, kriisistä toipuva ja vakaa. Tulosten perusteella SOM-malli on parempi tai vähintään yhtä hyvä kuin perinteinen logit-malli luokittelemaan otosaineisto ja ennustamaan vuonna 2007 alkanut globaali finanssikriisi. Mallin toimivuutta testataan muuttamalla mallin oletuksia, politiikantekijän preferenssejä ja ennustushorisonttia.

JEL-koodit: E44, E58, F01, F37, G01.
Asiasanat: systeemiset finanssikriisit, systeemiriski, SOM, visualisointi, ennustaminen, makrovakauden valvonta

Non-technical summary

The recent global financial crisis has demonstrated the importance of understanding sources of domestic and global vulnerabilities that may lead to a systemic financial crisis. Early identification of sources of vulnerability is important as it would allow introduction of policy actions to decrease further build up of vulnerabilities or enhance the shock absorption capacity of the financial system.

Much of the empirical literature deals with early warning systems (EWSs) that rely on conventional statistical modelling methods, such as the univariate 'signals' approach or multivariate logit/probit models. Given the changing nature of financial crises, stand-alone numerical predictions are unlikely to be able to thoroughly describe them. As a complement to numerical predictions, this motivates the development of tools with clear visual capabilities, enabling real human perception.

Dimensionality of the problem complicates visualization, since a large number of indicators are often required to accurately assess vulnerabilities to a financial crisis. In addition to the limitation of standard two- and three-dimensional visualizations in describing higher dimensions, there are challenges of including a temporal or cross-sectional dimension. Moreover, while composite indices of leading indicators and predicted probabilities of EWSs enable comparison across countries and over time, such indices fall short in representing sub-dimensions of the problem. Methods for exploratory data analysis can to some extent overcome these types of shortcomings. Exploratory data analysis attempts to describe the phenomena of interest in easily understandable forms by illustrating the structures in data. The Self-Organizing Map (SOM) is a method that combines the aims of projection and clustering techniques. It can provide an easily interpretable non-linear description of the multidimensional data distribution on a two-dimensional plane without losing sight of individual indicators. Thus, the two-dimensional output of the SOM makes it particularly useful for static visualizations, or summarizations, of large amounts of information.

This paper describes a methodology to map the state of financial stability and the sources of systemic risks on a two-dimensional plane. The Self-Organizing Finan-

cial Stability Map (SOFSM) enables a two-dimensional representation of a multidimensional financial stability space and allows disentangling the individual sources of vulnerabilities impacting on systemic risks. The map can be used to monitor macro-financial vulnerabilities by locating a particular country in the financial stability cycle: being it either in the pre-crisis, crisis, post-crisis or tranquil state. In addition, the SOM model performs as well or better than a logit model in classifying in-sample data and predicting out-of-sample the global financial crisis that started in 2007. Robustness of the SOFSM is tested by varying the thresholds of the models, policymaker preferences, and the forecasting horizon.

1 Introduction

The recent global financial crisis has demonstrated the importance of understanding sources of domestic and global vulnerabilities that may lead to a systemic financial crisis.[3] Early identification of financial stress would allow policymakers to introduce policy actions to decrease or prevent further build up of vulnerabilities or otherwise enhance the shock absorption capacity of the financial system. Finding the individual sources of vulnerability and risk is of central importance since that allows targeted actions for repairing specific cracks in the financial system.

Much of the empirical literature deals with early warning systems (EWSs) that rely on conventional statistical modelling methods, such as the univariate signals approach or multivariate logit/probit models.[4] However, financial crises are complex events driven by non-linearly related and non-normally distributed economic and financial factors.[5] These non-linearities derive, for example, from the fact that crises become more likely as the number of fragilities increase. Due to distributional assumptions, conventional statistical techniques may fail in modelling these events. Novel EWSs attempt to model these complex relationships by applying non-linear techniques (Demyanyk and Hasan, 2010). For example, Peltonen (2006) and Fioramanti (2008) show that a neural network outperforms a probit model in predicting currency and debt crises. However, while the utilization of non-linear techniques may increase *a posteriori* prediction accuracies to a minor extent, Peltonen (2006) and Berg *et al.* (2005) demonstrate that the results of *a priori* predictions of financial crises remain disappointing. Given the changing nature of the occurrences of these extreme events, stand-alone numerical analyzes are unlikely to comprehensively describe them. As a complement, this motivates the development of tools with clear visual capabilities and intuitive interpretability, enabling real human perception.

[3] Cardarelli *et al.* (2011) show that out of 113 financial stress episodes for 17 key advanced economies, 29 were followed by an economic slowdown and an equal number by recessions.
[4] Logit and probit models have been applied frequently to predicting financial crises. For example, Berg and Pattillo (1999) apply a discrete choice model to predicting currency crises; Fuertes and Kalotychou (2006) to predicting debt crises; and Lo Duca and Peltonen (2011) to predicting systemic crises. An exception is the univariate non-parametric indicator proposed by Kaminsky *et al.* (1998), and its subsequent versions. See Berg *et al.* (2005) for a comprehensive review.
[5] Fioramanti (2008), Sarlin and Marghescu (2009) and Lo Duca and Peltonen (2011) show that indicators of debt, currency, and systemic crises are non-linearly related.

One reason interpretability of the monitoring systems has not been adequately addressed is the complexity of the problem. A large number of indicators are often required to accurately assess the sources of financial instability. As with raw statistical tables, standard two- and three-dimensional visualizations have, of course, their limitations for high dimensions, not to mention the challenge of including a temporal or cross-sectional dimension or assessing multiple countries over time. Although composite indices of leading indicators and predicted probabilities of EWSs enable comparison across countries and over time, these indices fall short in disentangling the sources of vulnerability.[6] The recent work by IMF staff on the Global Financial Stability Map (GFSM) (Dattels *et al.*, 2010) has sought to overcome this challenge by a mapping of six composite indices.[7] Even here, however, the GFSM spider chart visualization of six indices falls short in disentangling individual sources. Familiar limitations of spider charts are, for example, the facts that area scales as the square of values, while the area itself depends on the order of dimensions. In addition, the use of adjustment based on market and domain intelligence, especially during crisis episodes, and the absence of a systematic evaluation gives neither a transparent data-driven measure of financial stress nor an objective anticipation of the GFSM's future precision. Indeed, the GFSM comes with the following caveat: "*given the degree of ambiguity and arbitrariness of this exercise the results should be viewed merely illustrative*".[8]

Methods for exploratory data analysis such as projection and clustering techniques may help in overcoming these shortcomings by illustrating data structures in easily understandable forms. The Self-Organizing Map (SOM) (Kohonen, 1982; 2001) is a method that combines the aims of projection and clustering techniques. It is capable of providing an easily interpretable non-linear description of the multidimen-

[6] There exist several composite indices for measuring financial tensions, e.g. Illing and Liu (2006), Cardarelli *et al.* (2011) and Lo Duca and Peltonen (2011). These will be further discussed in Section 2.
[7] The GFSM has appeared quarterly in the Global Financial Stability Report (GFSR) since April 2007.
[8] The authors state that the definitions of starting and ending dates of the assessed crisis episodes are arbitrary. Similarly, the assessed crisis episodes are arbitrary, as some episodes in between the assessed ones are disregarded, such as Russia's default in 1999 and the collapse of Long-Term Capital Management. Introduction of judgment based on market intelligence and technical adjustments are motivated when the GFSM is "*unable to fully account for extreme events surpassing historical experience*", which is indeed an obstacle for empirical models, but also a factor of uncertainty in terms of future performance since nothing assures manual detection of vulnerabilities, risks and triggers.

sional data distribution on a two-dimensional plane without losing sight of individual indicators. The two-dimensional output of the SOM makes it particularly useful for static visualizations, or summarizations, of large amounts of information (Back *et al.*, 1998).

By 2005, over 7700 works had featured the SOM (Pöllä *et al.*, 2009). While extensively applied to topics in engineering and medicine, the literature is short of thorough testing of the SOM for financial stability monitoring. In the emerging market context, Arciniegas and Arciniegas Rueda (2009), Sarlin (2011), Sarlin and Marghescu (2011) and Resta (2009) have applied the SOM to indicators of currency crises, debt crises and general economic and financial performance, respectively. The SOM has not, to our knowledge, been earlier applied to monitoring systemic risk or assessing the global dimensions of financial stability, including global macro-financial proxies as well as individual advanced and emerging market economies. Indeed, of the above applications, only Sarlin and Marghescu (2011) perform a thorough, systematic evaluation of the model's predictive capabilities.

The main contribution of this paper is to lay out a methodology for mapping the state of financial stability on a two-dimensional plane. As an enhancement to the GFSM approach, the Self-Organizing Financial Stability Map (SOFSM) not only allows disentangling the individual sources of vulnerability, but also performs well as an EWS in predicting out-of-sample systemic financial crises. The SOM parameter values for the final model are chosen based on a training framework aiming at a parsimonious, objective and interpretable model. Robustness of the SOFSM is tested by varying the thresholds of the models, policymaker preferences, and the forecasting horizon. In addition, when assessing a topologically ordered SOFSM, the concept of a financial stability neighborhood represents contagion of instabilities through similarities in the current macro-financial conditions. That is, a crisis in one position on the map indicates propagation of financial distress to adjacent locations. This type of representation may help in identifying the changing nature of crises. Further, inspired by Minsky's (1982) and Kindleberger's (1996) vindicated financial fragility view of a credit or asset cycle, we introduce the notion of the financial stability cycle. We show how the SOFSM can be used to monitor macro-financial vulnerabilities by locating a

country in the financial stability cycle: being it either in the pre-crisis, crisis, post-crisis or tranquil state. We visualize samples of the panel dataset, cross-sectional and temporal data, on the two-dimensional map, and compute and visualize aggregates for the world, emerging market economies and advanced economies. The SOFSM enables disentangling the specific threats, risks and triggers, and should be treated as a starting point rather than an ending point for financial stability analysis.

The paper is structured as follows. Section 2 introduces the SOM, the data and the evaluation framework. We present the training process and evaluation of the SOFSM in Section 3, and provide visual analyzes in Section 4. Section 5 concludes with a presentation of our key findings and suggestions for future work.

2 Methodology

Methods for exploratory data analysis fall, in general, into two groups: data and dimensionality reduction methods. Clustering methods attempt to reduce the amount of data by enabling analysis of a few mean profiles (clusters), but do not seek to project data to an easily interpretable format. Dimensionality reduction methods, e.g. Sammon's (1969) mapping and its variants (Cox and Cox, 2001), project high-dimensional data onto a lower dimension, while attempting to preserve the structure of the dataset. Unlike clustering methods, however, projection methods do not generally seek to reduce or distil the amount of presented data. The SOM combines the objectives of projection and clustering techniques.

2.1 Self-Organizing Maps (SOMs)

The SOM is a projection and clustering tool that uses an unsupervised learning method developed by Kohonen (1982). It differs from projection techniques like multidimensional scaling by performing a mapping from the input data space Ω onto a k-dimensional array of output nodes instead of into a continuous space and by attempting to preserve the neighbourhood relations in data rather than absolute distances. The vector quantization capability of the SOM allows modelling from the continuous space Ω, with a probability density function $f(x)$, to the grid of nodes, whose location

depend on the neighbourhood structure of the data Ω. On a two-dimensional grid, for example, the numbers on the *x*- and *y*-axes do not carry a numeric meaning in a parametric sense; they represent positions in the data space of the map, where each of these positions (*x*,*y*) is a mean profile. A second-level clustering can be applied on the nodes of the SOM, i.e. separation of data into nodes and nodes into clusters, as proposed in Vesanto and Alhoniemi (2000). They show that, compared to other clustering methods, the two-level SOM enhances the clustering through greater robustness on non-normally distributed data and the dual advantage of efficiency and speed. In Marghescu (2007), the data visualization features of the two-level SOM have been reviewed as better than those of other techniques. Information products of two-level SOMs have also been evaluated as superior than currently used methods by end-users within the domain of financial analysis (Eklund *et al.*, 2008). The brief description of the basic SOM algorithm follows that in Sarlin (2011). For further details on the SOM, see Kohonen (2001).

This study uses the Viscovery SOMine 5.1 package. [9] In addition to an easily interpretable visual representation, it employs the batch training algorithm, and thus processes data in batches instead of sequences. The most important advantage of the batch algorithm is the reduction of computational cost. The training process starts with initialization of the reference vectors set to the direction of the two principal components of the input data. The principal component initialization not only further reduces computational cost and enables reproducible results, but is also shown to be important for convergence when using the batch SOM (Forte *et al.*, 2002). Following the initialization, the batch training algorithm operates a specified number of iterations in two steps.

In the first step, each input data vector x is assigned to the best-matching unit (BMU) m_c:

$$\|x - m_c\| = \min_i \|x - m_i\| . \quad (1)$$

[9] There are several other implementations of the SOM. The seminal packages – SOM_PAK, SOM Toolbox for Matlab, Nenet, etc – are not regularly updated or adapted to their environment. Out of the newer implementations, Viscovery SOMine provides the needed techniques for interactive exploratory analysis (Moehrmann *et al.*, 2011). For a thorough discussion of SOM software and the implementation in Viscovery SOMine, see Deboeck (1998a; 1998b).

We employ a semi-supervised version of the SOM by also including class information when determining the BMU. In the second step, each reference vector m_i (where i=1,2,…,M) is adjusted using the batch update formula:

$$m_i(t+1) = .\frac{\sum_{j=1}^{N} h_{ic(j)}(t)x_j}{\sum_{j=1}^{N} h_{ic(j)}(t)} \quad (2)$$

where index j indicates the input data vectors that belong to node c, and N is the number of the data vectors. The neighbourhood function $h_{ic(j)} \in (0,1]$ is defined as the following Gaussian function:

$$h_{ic(j)} = \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right), \quad (3)$$

where $\|r_c - r_i\|^2$ is the squared Euclidean distance between the coordinates of the reference vectors m_c and m_i on the two-dimensional grid, and the radius of the neighbourhood $\sigma(t)$ is a monotonically decreasing function of time t. The radius of the neighbourhood begins as half the diagonal of the grid size ($\sigma = (X^2 + Y^2)^2 / 2$), and goes monotonically towards the specified tension value $\sigma(t) \in (0,2]$. The other parameters are map size (the number of nodes), map format (the ratio of X and Y dimensions), and the length of training (training cycles). Second-level clustering is done using an agglomerative hierarchical clustering. The following modified Ward's (1962) criterion is used as a basis for measuring the distance between two candidate clusters:

$$d_{kl} = \begin{cases} \frac{n_k n_l}{n_k + n_l} \cdot \|c_k - c_l\|^2 & \text{if } k \text{ and } l \text{ are adjacent} \\ \infty & \text{otherwise} \end{cases}, \quad (4)$$

where k and l represent two clusters, n_k and n_l the number of data points in the clusters k and l, and $\|c_k - c_l\|^2$ the squared Euclidean distance between the cluster

centres of clusters *k* and *l*. The Ward clustering is modified only to merge clusters with other topologically neighbouring clusters by defining the distance between non-adjacent clusters as infinitely large. The algorithm starts with each node as its own cluster and merges nodes for all possible numbers of clusters using the minimum Ward distance (1,2,…,*M*).

For the purpose of this analysis, the output of the SOM algorithm is visualized on a two-dimensional plane. The rationale for not using a one-dimensional map is the differences within clusters. A three-dimensional map, while adding a further dimension, impairs the interpretability of data visualizations. Here, the multidimensional space of the grid is visualized through layers, or "feature planes". For each individual indicator, a feature plane represents the distribution of its values on the two-dimensional map. As the feature planes are different views of the same map, one unique point represents the same node on all planes. We produce the feature planes here in colour. Cold to warm colours represent low to high values according to a colour scale below each feature plane. Shading on the two-dimensional map indicates the distance between each node and its corresponding second-level cluster centre, i.e. those close to the centre have a lighter shade and those farther away have a darker shade.

The quality of the map is usually measured in terms of quantization error, distortion measure and topographic error (see e.g. Vesanto *et al.*, 2003). As we have class information, we mainly use classification performance measures for evaluating the quality of the map.

2.2 Data

The indicators used in this paper are a replica of those in Lo Duca and Peltonen (2011). The dataset consists also of a database of systemic events and a set of vulnerability indicators commonly used in the macroprudential literature to predict financial crises. The quarterly dataset consists of 28 countries (10 advanced and 18 emerging economies) for the period 1990:1–2010:3. The data are retrieved from Haver Analytics, Bloomberg and Datastream.

Following Lo Duca and Peltonen (2011), the identification of systemic financial crises is performed using a Financial Stress Index (FSI). This approach provides an objective criterion for the definition of the starting date of a systemic financial crisis.[10] The idea behind the FSI is that the larger and broader the shock is (i.e. the more systemic the shock), the higher the co-movement among variables reflecting tensions in different market segments. By aggregating variables that measure stress across market segments, the FSI captures the starting and ending points of a financial crisis. The FSI is a country-specific composite index that covers the main segments of the domestic financial market: (1) the spread of the 3-month interbank rate over the 3-month government bill rate (Ind_1); (2) negative quarterly equity returns (Ind_2); (3) the realized volatility of the main equity index (as average daily absolute changes over a quarter) (Ind_3); (4) the realized volatility of the nominal effective exchange rate (Ind_4); and (5) the realized volatility of the yield on the 3-month government bill (Ind_5).[11] Each component *j* of the index for country *i* at quarter *t* is transformed into an integer from 0 to 3 according to the country-specific quartile of the distribution. For example, a value for component *j* falling into the third quartile of the distribution would be transformed to a "2". The FSI is computed for country *i* at time *t* as a simple average of the transformed variables as follows:

[10] There are several composite indices for measuring financial tensions. For example, Illing and Liu (2006) and Hakkio and Keeton (2009). Cardarelli *et al.* (2011) and Balakrishnan *et al.* (2009) constructed financial stability indices for a broad set of advanced and emerging economies, while Fidora and Straub (2009) created an index for the global economy.

[11] When the 3-month government bill rate is not available, the spread between interbank and T-bill rates of the closest maturity is used. The equity returns are multiplied by minus one, so that negative returns increase stress, while positive returns are set to 0. When computing realized volatilities for components Ind_{3-5}, average daily absolute changes over a quarter are used.

$$FSI_{i,t} = \frac{\sum_{j=1}^{5} q_{j,i,t}(Ind_{j,i,t})}{5} \quad (5)$$

To define systemic financial crises, the FSI is first transformed into a binary variable. We focus on episodes of extreme financial stress that have led in the past (on average) to negative consequences for the real economy in order to capture the systemic nature of the financial stress episodes. In practice, the binary variable takes a value 1 in the quarter when the FSI moves above the predefined threshold of the 90th percentile of the country distribution. This approach identifies a set of 94 systemic events. Next, we set the class variable C18 to 1 in the six quarters preceding the systemic financial crisis, and to 0 in all other periods, and define this as the "pre-crisis" period.[12] The "pre-crisis" class variable mimics an ideal leading indicator that perfectly signals a systemic financial crisis in the six quarters before the event.

To analyze the sources of systemic risk and vulnerability, we use the same indicators as in Lo Duca and Peltonen (2011). The set of indicators consists of commonly used metrics in the macroprudential literature for capturing the build-up of vulnerabilities and imbalances in the domestic and global economy (e.g. Borio and Lowe, 2002; 2004; Alessi and Detken, 2011). Our key variables are asset price developments and valuations, and variables proxying for credit developments and leverage. In addition, traditional variables (e.g. budget deficit and current account deficit) are used to control for vulnerability stemming from macroeconomic imbalances.

Following the literature, we construct several transformations of the indicators (e.g. annual changes and deviations from moving averages or trends) to proxy for misalignments and a build up of vulnerabilities. To proxy for global macro-financial imbalances and vulnerabilities, we calculate a set of global indicators by averaging the transformed variables for the United States, the Euro area, Japan and the United Kingdom.[13] The final set of indicators are chosen based on their univariate performance in predicting systemic events and are shown in Figure 1.

[12] In addition to C18, which is predicted with the benchmark model, we set the class variables C24, C12 and C6 to 1 in the 8, 4 and 2 quarters before the systemic event. C0 is the crisis dummy, while P6, P12, P18 and P24 are set to 1 in the 2, 4, 6 and 8 quarters after the systemic event. These are used to assess the financial stability cycle.

[13] Qualitatively similar results are obtained when global variables are constructed as simple averages of variables of all countries in the sample.

Statistical properties of the chosen best-performing indicators (Table 1) reveal that the data are significantly skewed and non-mesokurtic, and thus do not exhibit normal distributions. To take into account cross-country differences and country-specific fixed effects, we follow Kaminsky *et al.* (1998) by measuring indicators in terms of country-specific percentiles. While such outlier trimming is unnecessary for the clustering of the SOM, an even distribution is highly desirable for visualization.

Finally, the analysis is conducted in a real-time fashion to the extent possible. Thus, we take into account publication lags by using lagged variables. For GDP, money and credit related indicators, the lag ranges from 1 to 2 quarters depending on the country. We also de-trend variables and measure indicators in terms of country-specific percentiles using the latest available information.

(INSERT TABLE 1 HERE)

2.3 Evaluation framework

To evaluate the performance of models in terms of predicting systemic financial crises, we compute a set of accuracy or goodness-of-fit measures. As we have class information, we use classification performance measures for finding the optimal model rather than the traditional SOM quality measures. We classify the outcomes into combinations of predicted and actual classes using a contingency matrix.

		Actual class	
		1	-1
Predicted class	1	*True positive (TP)*	*False positive (FP)*
	-1	*False negative (FN)*	*True negative (TN)*

Based on the elements of the matrix, we compute ratios for measuring performance: recall, precision, False Positive (FP), True Positive (TP), False Negative (FN) and True Negative (TN) rates, and overall accuracy.[14] Due to unbalanced class sizes and differences in class importance, the above measures are sometime unsuited

[14] Recall positives = TP/(TP+FN), Recall negatives = TN/(TN+FP), Precision positives = TP/(TP+FP), Precision negatives = TN/(TN+FN), Accuracy = (TP+TN)/(TP+TN+FP+FN), TP rate = TP / (TP + FN), FP rate = FP/(FP+TN), FN rate = FN/(FN+TP) and TN rate = TN/(FP+TN).

to summarizing evaluations of crisis predictions. By assigning every object to the tranquil class, we would achieve a useless classifier for policy action, but still a high proportion of correct predictions (80%). This motivates using a common measure in information retrieval for evaluating performance on unbalanced class sizes. Matthews Correlation Coefficient (MCC) (Matthews, 1975) measures the correlation between the actual and predicted classes. It is defined in the range [-1,1], where -1 represents an inverse prediction and 1 a perfect prediction.[15] The costs of FNs and FPs might be asymmetric, where the weight depends on the policymaker's preferences between giving false signals of crisis and tranquil periods. To calibrate an optimal model and threshold for policy action, we adapt the approach pioneered in Demirgüç-Kunt and Detragiache (2000) with the technical implementation suggested by Alessi and Detken (2011). The loss function of the policymaker is thus defined as:

$$L(\mu) = \mu(FN/(FN+TP)) + (1-\mu)(FP/(FP+TN)), \quad (6)$$

where the parameter μ represents the relative preference of the policymaker between FNs and FPs. When $\mu = 0.5$, the policymaker is equally concerned about missing crises and issuing false signals. She is less concerned about issuing false alarms when $\mu > 0.5$ and more concerned when $\mu < 0.5$. To find out the usefulness of our predictions, we subtract the loss from the best-guess of the policy maker. This is given by $Min(\mu, 1-\mu)$, i.e., the expected value of a guess with the given preferences. From this, we obtain the usefulness of the model:

$$U = Min(\mu, 1-\mu) - L(\mu). \quad (7)$$

We do not explicitly assess the extent to which policymakers might be more or less concerned about failing to identify an impending crisis than issuing a false alarm. Missing a crisis may often, however, be more expensive than an internal alarm for further in-depth investigation of the vulnerabilities and risks. On the other hand, given the risks associated with self-fulfilling prophecies, a publicly reported false alarm can have costs on par with failure to not identify a crisis. In this paper, we use

[15] The MCC is computed as follows: $MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$.

as a benchmark model a policymaker with $\mu = 0.5$, but test model robustness by varying the preferences. The preferences of 0.5 could be considered those of policymaker who does not want to make mistakes by being equally concerned about missing crises and issuing false alarms.

Using receiver operating characteristics (ROC) curves and the area under the ROC curve (AUC), we measure the global performance of the models. The ROC curve shows the trade-off between the benefits and costs of choosing a certain threshold. When two models are compared, the better model has a higher benefit (expressed in terms of TP rate on the vertical axis) at the same cost (expressed in terms of FP rate on the horizontal axis).[16] The size of the AUC is estimated using trapezoidal approximations. It measures the probability that a randomly chosen crisis observation is ranked higher than a tranquil one. A random ranking has an expected AUC of 0.5, while a perfect ranking has an AUC equal to 1.

3 Self-Organizing Financial Stability Map

3.1 Training the Self-Organizing Financial Stability Map

Analysis of the financial stability cycle is enabled by introducing class variables representing different time periods: pre-crisis (C24, C18, C12, C6), crisis (C0), post-crisis (P6, P12, P18, P24) and tranquil (T0) periods. The pre- and post-crisis periods range from 24 months before to 24 months after a crisis. In contrast to Sarlin and Marghescu (2011), where the classes are not used in determining the BMU (Eq. 1), but are used within the updates of the reference vectors (Eq. 2), we let them have a minor impact by giving them a weight of 0.20 when determining the BMUs. More precisely, the importance of class distance is set to 0.20 in Eq. 1. Though we give the classes only a minor weight, we employ a semi-supervised SOM. This has a cost of lower classification and prediction accuracy, but yields the benefit of easier interpreta-

[16] In general, the ROC curve plots, for the whole range of measures, the conditional probability of positives to the conditional probability of negatives: $ROC = \frac{P(x|positive)}{P(x|negative)}$.

tion of the financial stability cycle. To partition the map according to the stages in the financial stability cycle, the nodes of the map are clustered with respect to the class variables using Ward clustering. Our crisp clustering given by the lines that separate the map into four parts should only be interpreted as aid in finding the four stages of the financial stability cycle, not as completely distinct clusters.

We obtain the predictive feature of the model by assigning to each data point the C18 (as well as C6, C12 and C24 when testing robustness) value of a node into which the data point is mapped with Eq. 1. The performance of a model is evaluated using the above introduced usefulness for a policymaker. The performance is computed using static and pooled models, i.e. the coefficients or maps are not re-estimated recursively over time and across countries. Following Fuertes and Kalotychou (2006), it can be assumed that by not varying the specification over time or across countries, the parsimonious models better generalize in-sample data and predict out-of-sample data. Although static models have the drawback of ignoring the latest available information, they allow for more thorough evaluation and comparison of model performance as well as better generalization. While Peltonen and Lo Duca (2011) include interaction terms, we do not replicate them since they are included in the SOFSM processing *per se*. To test the predictability of the 2008–2009 financial crisis, the training set spans 1990:4–2005:1, while the test set spans 2005:2–2009:2. The training framework and choice of map is implemented with respect to three aspects: (1) the model does not overfit the in-sample data (parsimonious); (2) the framework does not include out-of-sample performance (objective); and (3) visualization is taken into account (interpretability). For a parsimonious, objective and interpretable model, we employ the following training framework.

1. Train and evaluate in terms of in-sample usefulness models for $\sigma \in \{0.0001, 0.3, 0.5, 0.75, 1.0, 1.5, 2.0\}$ and $M \in \{50, 100, 150, 200, 250, 300, 400, 500, 600, 1000\}$.[17]

[17] We keep constant the map format (75:100) and the training length. Kohonen (2001) recommends that the map ought be oblong rather than square. To have a comparable training length for different parameters, we use an implementation in SOMine with an increasing function of map size and decreasing of data points, among other things. The varied parameters, M and tension σ, have the following effect on performance: an increase in the M value increases the in-sample usefulness, where $U \to 0.5$ when $M \to \infty$, but decreases out-of-sample usefulness. Increases in tension decrease quantization accuracy, and thus in-sample usefulness, but do not have a direct effect on out-of-sample performance. In fact, if

For each model, set the threshold such that the usefulness is maximized. For each *M*-value, order the models in a descending order.

2. Find for each *M*-value the first model with usefulness equal to or better than that of the logit model. Choose none of the models if for an *M*-value all (or none) of the models' usefulness exceed that of the logit model.

3. Evaluate the interpretability of the models chosen in Step 2. Choose the one that is most interpretable.

The evaluation results are shown in Table 2. For $M \in \{50,400,500,600,1000\}$ no model is chosen for analysis, as they never or always exceed the usefulness of the logit model (*U*=0.25). Finally, of the five highlighted models, we select the one with *M*=150 and $\sigma = 0.5$ (shown in bold) for its interpretability.

(INSERT TABLE 2 HERE)

The chosen map has 137 nodes on an 11x13 grid and is trained with a tension of 0.5. Figures 1–3 present the two-dimensional SOM grid, the feature planes for the 14 indicators and the feature planes for the class variables. The feature planes in Figure 3 show the real distribution of the classes on the map, while the lines that split the maps into four parts show crisp clustering of the nodes based on all class variables (except PPC0). The feature plane PPC0, with a high frequency on the border between the post- and pre-crisis cluster, represents the simultaneous occurrence of a pre- and post-crisis period. In this case, the cycle need not include the tranquil stage if a new pre-crisis period is entered directly after the previous event. Using the distribution of the class variables, the four clusters are labelled according to the stages of the financial stability cycle. The upper left cluster represents the pre-crisis cluster (Pre crisis), the lower left represents the crisis cluster (Crisis), the centre and lower-right cluster represents the post-crisis cluster (Post crisis) and the upper right represents the tranquil cluster (Tranquil). The main characteristics of the clusters can be derived from the feature planes in Figures 2–3. In contrast to EWSs using supervised methods, such as discrete choice techniques, the SOM model enables simultaneous assessment of the

M equals the cardinality of *x*, then perfect in-sample performance may be obtained by each *M* attracting one data. This would, however, be an overfitted model for out-of-sample prediction.

correlations with all four stages of the financial stability cycle. Thus, new models need not be derived for different forecasting horizons or definitions of the dependent variable. By assessing the feature planes of the SOFSM, the following strong correlations are found, for example. The strongest early signs of a crisis are high domestic and global real equity growth and equity valuation, while most important late signs of a crisis are domestic and global real GDP growth, and domestic real credit growth, leverage, budget surplus, and CA deficit. The highest values of global leverage and real credit growth in the crisis cluster exemplify the fact that increases in some indicators may reflect a rise in financial stress only up to a specific threshold. Increases beyond that level are, in this case, more concurrent than preceding signals of a crisis. Similarly, budget deficits characterize the late post-crisis and early tranquil periods, while surpluses are signals of impending instabilities. The characteristics of the financial stability states are summarized in Table 3.

(INSERT TABLE 3 HERE)

The topological ordering of the SOFSM enables assessing, in terms of macro-financial conditions, neighbouring financial states of a particular position on the map. Transmission of financial contagion is often defined by other types of neighbourhood measures such as financial or trade linkages, proxies of financial shock propagation, equity market co-movement or geographical relations (see for example Dornbusch *et al.* (2000) and Pericoli and Sbracia (2003)). When assessing the SOFSM, the concept of neighborhood of a country represents the similarity of the current macro-financial conditions. Thus, a crisis in one position on the map indicates propagation of financial instabilities to adjacent locations. This type of representation may help in identifying events surpassing historical experience and the changing nature of crises.

(INSERT FIGURE 1–3 HERE)

3.2 A Comparison with a Logit Model

The logit model is estimated using the same in-sample data as was used for the SOM models. The estimates are reported in Table 4 and are later used for predicting out-of-

sample data, while the in-sample and out-of-sample performance for the benchmark models are shown in Table 5. As is shown in Figure 4, for maximizing the usefulness for a policymaker the map is classified into two parts, where the shaded area represents early warning nodes and the rest tranquil nodes. For the benchmark models ($\mu = 0.5$ and C18), the overall performance is similar between the SOM model and the logit model. On the train set, the SOM model performs slightly better than the logit model in terms of usefulness, recall positives, precision negatives and the AUC measure, while the logit model outperforms on the other measures. The classification of the models are of opposite nature, as the SOM issues more false alarms (FP rate=31%) than it misses crises (FN rate=19%), whereas the logit model misses more crises (31%) than it issues false alarms (19%). That explains also the difference in the overall accuracy, since the class sizes are unbalanced (20% crisis and 80% tranquil periods). The performance of the models on the test set differs, in general, similarly as the performance on the train set, except for the SOM having slightly higher overall accuracy. This is, in general, due to the higher share of crisis episodes in the out-of-sample dataset.

We test the robustness of the SOFSM with respect to policymaker's preferences ($\mu = 0.4$ and $\mu = 0.6$), forecasting horizon (6, 12 and 24 months before a crisis) and thresholds. Changes in the policymaker's preferences affect the number of early warning nodes, as is shown in Figure 4. The results of the robustness tests are shown in Tables 6–7 and Figure 5. Table 6 shows the performance over different policymaker's preferences, Table 7 over different forecasting horizons and Figure 5 and Tables 6–7 over all possible thresholds. For a policymaker less concerned about issuing false alarms ($\mu = 0.6$), the performance of the models are similar, except for higher usefulness of the SOM model. This confirms that the SOM better detects the rare crisis occurrences. For a policymaker less concerned about missing crises ($\mu = 0.4$), the usefulness of the models is similar, but the nature of the prediction is reversed; the SOM issues less false alarms than it misses crises, whereas the logit model issues more false alarms than misses crises. Over different forecasting horizons, the in-sample performance is generally similar. However, the out-of-sample usefulness, with the exception of C12, is better for the SOM model than for the logit model. Interestingly, the logit model fails to yield any usefulness (U=0.02) at a fore-

casting horizon of 6 months. Finally, the AUC measure summarizes the performance over all possible thresholds using the plots shown for the benchmark models in Figure 5. It is the only measure to consistently show superior performance for the SOM model. A caution regarding the AUC measure is, however, that parts of the ROC curve that are not policy relevant are included in the computed area. When comparing usefulness for each pair of models, the SOM model shows consistently equal or superior performance except for a single out-of-sample evaluation with a forecasting horizon of 12 months. To sum up, the SOM performs, in general, as well or better than a logit model in both classifying the in-sample data and in predicting the global financial crisis that started in 2007.

(INSERT TABLE 4–7 HERE)
(INSERT FIGURE 4–5 HERE)

4 Mapping the State of Financial Stability

In this section, we use the SOFSM for mapping macro-financial conditions and the state of financial stability. We map samples of the panel dataset by showing cross-sectional and temporal data on the two-dimensional SOM grid. We also compute aggregates for groups of countries for exploring states of financial stability globally, in advanced countries and in emerging economies. Eq. 1 is used for mapping data points onto the grid, i.e. they are mapped to their BMU. Consecutive time-series data are linked with lines.

4.1 Cross-sectional and temporal analysis on the SOFSM

For a simultaneous temporal and comparative analysis, we map the evolution of macro-financial conditions for the United States and the Euro area in Figure 6. The data for both "countries" represent the first quarters of 2002 to 2010. Without a precise empirical treatment for accuracy, the map well recognizes for both countries the pre-crisis, crisis and post-crisis stages of the financial stability cycle by circulating

around the map during the analyzed period. Interestingly, the Euro area is located in the tranquil cluster in 2010Q1 (as well as in Q3 as is shown in Figure 7). This indicates that the aggregated macroprudential metrics for the Euro area as a whole do not reflect the ongoing fiscal or banking crises in the Euro area periphery. It also coincides with a relatively low FSI for the aggregate Euro area. This can be explained by the weaknesses and financial stress in smaller economies being averaged out by improved macro-financial conditions in larger Euro area economies, highlighting the importance of country-level analysis.

Figure 7 represents a cross-section mapping of all countries in 2010Q3. The countries are divided into three groups of financial stability states. The map indicates risks in several emerging market economies (Mexico, Turkey, Argentina, Brazil, Taiwan, Malaysia and the Philippines), while most of the advanced economies are in the lower right corner of the map (post-crisis and tranquil cluster). Three countries (Singapore, South Africa and India) sit on the border of the tranquil and pre-crisis clusters, which is an indication of a possible future transition to the pre-crisis cluster. For this type of cross-sectional data, the topological ordering of the SOFSM enables assessing propagation of financial instabilities to adjacent macro-financial locations. When the SOFSM does not account for events surpassing historical data, as empirical models of non-stationary processes may do, this type of representation may help in identifying the changing nature of crises. For this cross section (Figure 7), a crisis in, say, Argentina and Brazil would as well indicate possible financial distress in neighbouring countries (Taiwan, Mexico and Turkey).

(INSERT FIGURE 6–7 HERE)

4.2 Exploring aggregate financial stability states on the SOFSM

In this section, we map data as above for just three country aggregates: the world, emerging market economies and advanced economies. We compute the state of financial stability for the aggregates by weighting the indicators for the countries in our

sample using stock market capitalization.[18] These aggregates can, like any data point, be projected onto the map using Eq. 1. Figure 8 shows the evolution of global macro-financial conditions in the first quarters of 2002 to 2010. The global state of financial stability enters the pre-crisis cluster in 2006Q1 and the crisis cluster in 2007Q1. It moves via the post-crisis cluster to the tranquil cluster in 2010Q1. This coincides with the global evolution of the FSI. The separation of the global aggregate into emerging market and advanced economies is shown in Figure 9. The mapping of the advanced economy aggregate is a copy of the world aggregate, which is mainly a result of the small share of stock market capitalization in the emerging world. Notably, the emerging market movements are also similar to those in advanced economies. While the emerging market cycle moves around that of the advanced economies, it does not indicate significant differences in the timeline or strength of financial stress.

(INSERT FIGURE 8–9 HERE)

5 Conclusions

This paper creates a Self-Organizing Financial Stability Map for visualizing the sources of systemic risks on a two-dimensional plane and for predicting systemic financial crises. The SOFSM is a two-dimensional representation of a multidimensional financial stability space that allows disentangling the individual sources of vulnerabilities impacting on systemic risks. In addition, the model can be used to monitor macro-financial vulnerabilities by locating a country in the financial stability cycle: being it either in the pre-crisis, crisis, post-crisis or tranquil state. Our results indicate the SOM model performs as well or better than a logit model in classifying in-sample data and predicting the global financial crisis that started in 2007. We test the robustness of the SOFSM by varying the thresholds of the models, the policymaker's preferences, and the forecasting horizon. Future work should focus on the visual representation of the SOFSM by including membership degrees to cluster centres and by an in-depth assessment of the financial stability cycle using transition probabilities between nodes and clusters.

[18] The advanced economies are Australia, Denmark, Euro area, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The emerging market economies are Argentina, Brazil, China, Czech Republic, Hong Kong, Hungary, India, Indonesia, Malaysia, Mexico, the Philippines, Poland, Russia, Singapore, South Africa, Taiwan, Thailand and Turkey.

References

Alessi, L., Detken, C., 2011. Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy* 27(3), 520–533.

Arciniegas Rueda, I.E., Arciniegas, F., 2009. SOM-based data analysis of speculative attacks' real effects. *Intelligent Data Analysis* 13(2), 261–300.

Back, B., Sere, K., Vanharanta, H., 1998. Managing Complexity in Large Data Bases using Self-Organizing Maps. *Accounting, Management and Information Technologies* 8(4), 191–210.

Balakrishnan, R., Danninger, S., Elekdag, S., Tytell, I., 2009. The Transmission of Financial Stress from Advanced to Emerging Economies. IMF Working Paper, WP/09/133.

Berg, A., Borensztein, E., Pattillo, C., 2005. Assessing early warning systems: How have they worked in practice?. *IMF Staff Papers* 52, 462–502.

Berg, A., Pattillo, C., 1999. Predicting currency crises – the indicators approach and an alternative. *Journal of International Money and Finance* 18, 561–586.

Borio, C., Lowe, P., 2002. Asset Prices, Financial and Monetary Stability: Exploring the Nexus. BIS Working Papers, No. 114.

Borio, C., Lowe, P., 2004. Securing Sustainable Price Stability: Should Credit Come Back from the Wilderness?. BIS Working Papers, No. 157.

Cardarelli, R., Elekdag, S., Lall, S., 2011. Financial stress and economic contractions. *Journal of Financial Stability* 7(2), 78–97.

Cox, T.F., Cox, M.A.A., 2001. *Multidimensional Scaling*. Chapman & Hall/CRC, Florida.

Dattels, P., McCaughrin, R., Miyajim, K., Puig, J., 2010. Can you Map Global Financial Stability?. IMF Working Paper, WP/10/145.

Deboeck, G., 1998a. Software Tools for Self-Organizing Map, in: Deboeck, G., Kohonen, T., (Eds.), *Visual Explorations in Finance with Self-Organizing Maps*, Springer-Verlag, Berlin, pp. 179–194.

Deboeck, G., 1998b. "Best practices in data mining using self-organizing maps, in: Deboeck, G., Kohonen, T., (Eds.), *Visual Explorations in Finance with Self-Organizing Maps*, Springer-Verlag, Berlin, pp. 201–229.

Demirgüç-Kunt, A., Detragiache, E., 2000. Monitoring Banking Sector Fragility. A Multivariate Logit. *World Bank Economic Review* 14(2), 287–307.

Demyanyk, Y.S., Hasan, I., 2010. Financial crises and bank failures: a review of prediction methods. *Omega* 38(5), 315–324.

Dornbusch, R., Park, Y.C., Claessens, S., 2000. Contagion: How it Spreads and How it can be Stopped. *World Bank Research Observer* 15, 177–197.

Eklund, T., Back, B., Vanharanta, H., Visa, A., 2000. Evaluating a SOM-based financial benchmarking tool. *Journal of Emerging Technologies in Accounting* 5(1), 109–127.

Illing, M., Liu, Y., 2006. Measuring financial stress in a developed country: An application to Canada. *Journal of Financial Stability* 2(3), 243–65.

Fioramanti, M., 2008. Predicting sovereign debt crises using artificial neural networks: a comparative approach. *Journal of Financial Stability* 4(2), 149–164.

Forte, J.C., Letrémy, P., Cottrell, M., 2002. Advantages and drawbacks of the Batch Kohonen algorithm, in: Verleysen, M., (Ed.), *Proceedings of the 10th European Symposium on Neural Networks*, Springer-Verlag, Berlin, pp. 223–230.

Fuertes, A.M., Kalotychou, E., 2006. Early Warning System for Sovereign Debt Crisis: the role of heterogeneity. *Computational Statistics and Data Analysis* 5, 1420–1441.

Hakkio, C.S., Keeton, W.R., 2009. Financial Stress: What is it, How can it be measured and Why does it matter?. Federal Reserve Bank of Kansas City Economic Review, Second Quarter 2009, 5–50.

Kaminsky, G., Lizondo, S., Reinhart, C., 1998. Leading Indicators of Currency Crises. *IMF Staff Papers* 45(1), 1–48.

Kindleberger, C., 1996. Maniacs, Panics, and Crashes. Cambridge University Press, Cambridge.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 66, 59–69.

Kohonen, T., 1982. *Self-Organizing Maps*, 3rd edition. Springer-Verlag, Berlin.

Lo Duca, M., Peltonen, T.A., 2011. Macro-Financial Vulnerabilities and Future Financial Stress – Assessing Systemic Risks and Predicting Systemic Events. ECB Working Paper, No. 1311.

Marghescu, D., 2007. Multidimensional Data Visualization Techniques for Exploring Financial Performance Data, in: *Proceedings of 13th Americas Conference on Information Systems*, Keystone, Colorado, USA.

Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) – Protein Structure* 405(2), 442–45.

Minsky, H., 1982. Can "it" Happen Again?: Essays on Instability and Finance. M.E. Sharpe, Armonk, N.Y.

Moehrmann, J., Burkovski, A., Baranovskiy, E., Heinze, G.A., Rapoport, A., Heideman, G., 2011. A Discussion on Visual Interactive Data Exploration Using Self-Organizing Maps, in: Laaksonen, J., Honkela, T., (Eds.), *Proceedings of the 8th International Workshop on Self-Organizing Maps*, Springer-Verlag, Berlin, pp. 178–187.

Peltonen, T.A., 2006. Are emerging market currency crises predictable? A test. ECB Working Paper, No. 571.

Pericoli, M., Sbracia, M., 2003. A Primer on Financial Contagion. *Journal of Economic Surveys* 17, 571–608.

Pöllä, M., Honkela, T., Kohonen, T., 2009. Bibliography of Self-Organizing Map (SOM) Papers: 2002-2005 Addendum. TKK Reports in Information and Computer Science, Helsinki University of Technology, Report TKK-ICS-R24.

Resta, M., 2009. Early Warning Systems: an approach via Self Organizing Maps with applications to emergent markets, in: Apolloni, B., Bassis, S., Marinaro, M. (Eds.), *Proceedings of the 18th Italian Workshop on Neural Networks*, IOS Press, Amsterdam, pp. 176–184.

Sammon Jr., J.W., 1969. A Non-Linear Mapping for Data Structure Analysis. *IEEE Transactions on Computers* 18(5), 401–409.

Sarlin, P., 2011. Sovereign Debt Monitor: A Visual Self-Organizing Maps Approach, in: *Proceedings of the IEEE Symposium on Computational Intelligence for Financial Engineering & Economics*, IEEE Press, Paris, pp. 357–364.

Sarlin, P., Marghescu, D., 2011. Visual Predictions of Currency Crises using Self-Organizing Maps. *Intelligent Systems in Accounting, Finance and Management* 18(1), 15–38.

Ward Jr., J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58, 236–244.

Vesanto, J., Alhoniemi, E., 2000. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks* 11(3), 586–600.

Vesanto, J., Sulkava, M., Hollmén, J., 2003. On the decomposition of the self-organizing map distortion measure, in: *Proceedings of the Workshop on Self-Organizing Maps (WSOM'03)*, Springer-Verlag, Berlin, pp. 11–16.

Table 1: Statistical properties of the dataset

Type	Variable	Abbreviation	Mean	SD	Min.	Max.	Skew.	Kurt.	KSL	AD
Domestic	Inflation[a]	Inflation	0.89	5.17	-10.15	42.53	4.80	26.72	0.29*	263.90*
Domestic	Real GDP[b]	Real GDP growth	3.73	3.76	-17.54	14.13	-0.86	3.16	0.06*	11.34*
Domestic	Real credit to private sector to GDP[b]	Real credit growth	234.07	4724.00	-69.42	101870.34	20.76	429.59	0.51*	Inf*
Domestic	Real equity prices[b]	Real equity growth	5.93	33.01	-84.40	257.04	0.99	4.31	0.05*	7.28*
Domestic	Credit to private sector to GDP[a]	Leverage	3.48	51.64	-62.78	1673.04	22.76	673.35	0.29*	Inf*
Domestic	Stock market capitalisation to GDP[a]	Equity valuation	3.90	28.32	-62.79	201.55	0.77	2.41	0.03*	3.86*
Domestic	Current account deficit to GDP[c]	CA deficit	-0.02	0.07	-0.27	0.10	-0.98	0.73	0.09*	33.12*
Domestic	Government deficit to GDP[c]	Government deficit	0.01	0.05	-0.19	0.22	-1.09	3.46	0.09*	35.90*
Global	Inflation[a]	Global inflation	0.03	0.64	-1.33	2.29	0.71	1.28	0.08*	12.12*
Global	Real GDP[b]	Global real GDP growth	1.84	1.59	-6.34	4.09	-3.02	11.74	0.20*	122.16*
Global	Real credit to private sector to GDP[b]	Global real credit growth	3.87	1.68	-0.23	7.20	-0.21	-0.31	0.07*	8.82*
Global	Real equity prices[b]	Global real equity growth	2.31	19.08	-40.62	37.77	-0.57	-0.68	0.15*	41.90*
Global	Credit to private sector to GDP[a]	Global leverage	1.15	2.79	-2.79	11.21	1.84	3.40	0.22*	105.26*
Global	Stock market capitalisation to GDP[a]	Global equity valuation	0.89	17.41	-40.54	27.46	-0.50	-0.43	0.09*	19.11*

Notes: Transformations: [a], deviation from trend; [b], annual change; [c], level. KSL: Lilliefors' adaption of the Kolmogorov-Smirnov normality test. AD: the standard Anderson-Darling normality test. Significance levels: 1%, *.

Table 2: Characteristics of the financial stability states

Variable	Pre crisis		Crisis		Post crisis		Tranquil	
	Centre	Range	Centre	Range	Centre	Range	Centre	Range
Inflation	0.49	[0.22,0.66]	0.55	[0.30,0.69]	0.59	[0.26,0.76]	0.37	[0.17,0.68]
Real GDP growth	0.67	[0.40,0.80]	0.48	[0.14,0.83]	0.34	[0.25,0.50]	0.53	[0.30,0.72]
Real credit growth	0.66	[0.28,0.85]	0.55	[0.35,0.82]	0.39	[0.18,0.68]	0.43	[0.21,0.75]
Real equity growth	0.68	[0.41,0.85]	0.28	[0.16,0.58]	0.39	[0.23,0.80]	0.61	[0.40,0.74]
Leverage	0.63	[0.31,0.80]	0.59	[0.37,0.81]	0.52	[0.23,0.83]	0.29	[0.18,0.51]
Equity valuation	0.73	[0.62,0.80]	0.55	[0.27,0.81]	0.33	[0.17,0.66]	0.45	[0.30,0.63]
CA deficit	0.58	[0.30,0.78]	0.54	[0.26,0.80]	0.48	[0.25,0.77]	0.41	[0.19,0.66]
Government deficit	0.38	[0.19,0.74]	0.45	[0.22,0.62]	0.53	[0.32,0.85]	0.61	[0.26,0.85]
Global inflation	0.33	[0.08,0.61]	0.61	[0.34,0.76]	0.46	[0.20,0.79]	0.63	[0.11,0.90]
Global real GDP growth	0.67	[0.54,0.74]	0.67	[0.30,0.86]	0.29	[0.13,0.69]	0.45	[0.13,0.71]
Global real credit growth	0.55	[0.28,0.77]	0.86	[0.61,0.92]	0.37	[0.16,0.67]	0.33	[0.15,0.52]
Global real equity growth	0.72	[0.47,0.80]	0.4	[0.23,0.63]	0.34	[0.11,0.79]	0.54	[0.20,0.73]
Global leverage	0.35	[0.18,0.60]	0.79	[0.57,0.91]	0.58	[0.17,0.77]	0.33	[0.16,0.73]
Global equity valuation	0.67	[0.48,0.82]	0.81	[0.54,0.91]	0.36	[0.14,0.76]	0.27	[0.19,0.55]

Notes: Columns represent characteristics (cluster centre and range) of the financial stability states on the SOFSM and rows represent indicators. Since data are transformed to country-specific percentiles, the summary statistics are comparable across indicators and clusters.

Table 3: The evaluation of the SOFSM over *M* and σ values (μ=0.5 and forecasting horizon 6 quarters)

σ (tension) / *M* (# nodes)	0.001	0.3	0.5	0.75	1	1.5	2
50 (52)	0.24	0.23	0.22	0.21	0.21	0.20	0.20
100 (85)	0.27	0.25	0.23	0.22	0.21	0.21	0.21
150 (137)	0.29	0.24	**0.25**	0.23	0.21	0.23	0.21
200 (188)	0.29	0.29	0.29	0.24	0.23	0.22	0.21
250 (247)	0.30	0.29	0.29	0.24	0.25	0.21	0.22
300 (331)	0.32	0.33	0.30	0.28	0.25	0.23	0.22
400 (408)	0.40	0.40	0.38	0.33	0.30	0.27	0.27
500 (493)	0.42	0.40	0.40	0.36	0.33	0.28	0.27
600 (609)	0.43	0.43	0.41	0.36	0.33	0.28	0.27
1000 (942)	0.46	0.46	0.44	0.41	0.36	0.31	0.30

Notes: First models to outperform the logit model (U=0.25) per M value are highlighted in gray and the chosen map is shown in bold. The real number of nodes is shown in parenthesis since fulfilling the map ratio (75:100) affects the number of nodes.

Table 4: The estimates of the logit model (μ=0.5 and forecasting horizon 6 quarters)

Variable	*Estimate*	*Error*	*Z*	*Sig.*	
Intercept	-6.744	0.612	-11.024	0.000	***
Inflation	-0.100	0.300	-0.334	0.738	
Real GDP growth	0.076	0.334	0.229	0.819	
Real credit growth	-0.001	0.001	-0.613	0.540	
Real equity growth	1.791	0.382	4.685	0.000	***
Leverage	0.003	0.001	3.204	0.001	***
Equity valuation	0.002	0.001	2.689	0.007	***
CA deficit	1.151	0.308	3.741	0.000	***
Government deficit	0.076	0.342	0.223	0.823	
Global inflation	0.207	0.341	0.608	0.543	
Global real GDP growth	1.156	0.419	2.761	0.006	***
Global real credit growth	0.685	0.381	1.799	0.072	*
Global real equity growth	0.832	0.419	1.985	0.047	**
Global leverage	0.712	0.427	1.668	0.095	*
Global equity valuation	0.959	0.472	2.029	0.042	**

Notes: Significance levels: 1%, ***; 5 %, **; 10 %, *.

Table 5: Performance of the benchmark models on in-sample and out-of-sample data (μ=0.5 and forecasting horizon 6 quarters).

							Positives		*Negatives*					
Model	***Data set***	***Threshold***	***TP***	***FP***	***TN***	***FN***	***Precision***	***Recall***	***Precision***	***Recall***	***Accuracy***	***U***	***AUC***	***MCC***
Logit	Train	0.72	162	**190**	**830**	73	**0.46**	0.69	0.92	**0.81**	**0.79**	**0.25**	0.81	**0.44**
SOM	Train	0.60	**190**	314	706	**45**	0.38	**0.81**	**0.94**	0.69	0.71	**0.25**	**0.83**	0.40
Logit	Test	0.72	77	**57**	**249**	93	**0.57**	0.45	0.73	**0.81**	0.68	0.13	0.72	0.28
SOM	Test	0.60	**112**	89	217	**58**	0.56	**0.66**	**0.79**	0.71	**0.69**	**0.18**	**0.75**	**0.36**

Notes: The table reports results for the logit and SOFSM on the train and test data sets and the optimal threshold. Thresholds are calculated for μ=0.5 and forecastin horizon 6 quarters. The Table also reports in columns the following measures to assess the performance of the models: TP = True positives, FP = False positives, TN= True negatives, FN = False negatives, Precision positives = TP/(TP+FP), Recall positives = TP/(TP+FN), Precision negatives = TN/(TN+FN), Recall negatives = TN/(TN+FP), Accuracy = (TP+TN)/(TP+TN+FP+FN), usefulness U (see formulae 6 and 7), AUC = area under the ROC curve (TP rate to FP rate, see Section 2 and Figure 4) and MCC = (TP*TN-FP*FN)/$\sqrt{((TP+FP)(TP+FN)(TN+FP)(TN+FN))}$.

Table 6: Robustness tests on in-sample and out-of-sample data for different μ values (forecasting horizon 6 quarters)

								Positives		*Negatives*					
Model	***Data set***	***μ***	***Threshold***	***TP***	***FP***	***TN***	***FN***	***Precision***	***Recall***	***Precision***	***Recall***	***Accuracy***	***U***	***AUC***	***MCC***
Logit	Train	0.4	0.72	**162**	190	830	**73**	0.46	**0.69**	**0.92**	0.81	0.79	**0.16**	0.81	**0.44**
SOM	Train	0.4	0.75	153	**166**	**854**	82	**0.48**	0.65	0.91	**0.84**	**0.80**	**0.16**	**0.83**	**0.44**
Logit	Train	0.5	0.72	162	**190**	**830**	73	**0.46**	0.69	0.92	**0.81**	**0.79**	**0.25**	0.81	**0.44**
SOM	Train	0.5	0.60	**190**	314	706	**45**	0.38	**0.81**	**0.94**	0.69	0.71	**0.25**	**0.83**	0.40
Logit	Train	0.6	0.54	197	**381**	**639**	38	**0.34**	0.84	0.94	**0.63**	**0.67**	0.15	0.81	0.36
SOM	Train	0.6	0.50	**214**	419	601	**21**	**0.34**	**0.91**	**0.97**	0.59	0.65	**0.18**	**0.83**	**0.39**
Logit	Test	0.4	0.72	**77**	57	249	**93**	0.57	**0.45**	**0.73**	0.81	**0.68**	**0.07**	0.72	**0.28**
SOM	Test	0.4	0.75	76	**56**	**250**	94	**0.58**	**0.45**	**0.73**	**0.82**	**0.68**	**0.07**	**0.75**	**0.28**
Logit	Test	0.5	0.72	77	**57**	**249**	93	**0.57**	0.45	0.73	**0.81**	0.68	0.13	0.72	0.28
SOM	Test	0.5	0.60	**112**	89	217	**58**	0.56	**0.66**	**0.79**	0.71	**0.69**	**0.18**	**0.75**	**0.36**
Logit	Test	0.6	0.54	110	**109**	**197**	60	0.50	0.65	0.77	**0.64**	0.64	0.05	0.72	0.28
SOM	Test	0.6	0.50	**134**	**109**	**197**	**36**	**0.55**	**0.79**	**0.85**	**0.64**	**0.70**	**0.13**	**0.75**	**0.41**

Notes: See notes for Table 5.

Table 7: Robustness tests on in-sample and out-of-sample data for different horizons (μ=0.5)

								Positives		*Negatives*					
Model	***Data set***	***Horizon***	***Threshold***	***TP***	***FP***	***TN***	***FN***	***Precision***	***Recall***	***Precision***	***Recall***	***Accuracy***	***U***	***AUC***	***MCC***
Logit	Train	C6	0.72	70	**282**	**882**	21	**0.20**	0.77	0.98	**0.76**	**0.76**	**0.26**	0.81	**0.30**
SOM	Train	C6	0.51	**88**	530	634	**3**	0.14	**0.97**	**1.00**	0.54	0.58	**0.26**	**0.83**	0.27
Logit	Train	C12	0.72	117	**235**	**855**	48	**0.33**	0.71	**0.95**	**0.78**	**0.77**	**0.25**	0.80	**0.37**
SOM	Train	C12	0.69	**123**	267	823	**42**	0.32	**0.75**	**0.95**	0.76	0.75	**0.25**	**0.84**	**0.37**
Logit	Train	C18	0.72	162	**190**	**830**	73	**0.46**	0.69	0.92	**0.81**	**0.79**	**0.25**	0.81	**0.44**
SOM	Train	C18	0.60	**190**	314	706	**45**	0.38	**0.81**	**0.94**	0.69	0.71	**0.25**	**0.83**	0.40
Logit	Train	C24	0.58	**242**	286	673	**54**	0.46	**0.82**	**0.93**	0.70	0.73	0.26	0.81	0.45
SOM	Train	C24	0.63	233	**241**	**718**	63	**0.49**	0.79	0.92	**0.75**	**0.76**	**0.27**	**0.85**	**0.47**
Logit	Test	C6	0.72	18	**116**	**302**	40	0.13	0.31	0.88	**0.72**	**0.67**	0.02	0.57	0.02
SOM	Test	C6	0.51	**47**	205	213	**11**	**0.19**	**0.81**	**0.95**	0.51	0.55	**0.16**	**0.65**	**0.21**
Logit	Test	C12	0.72	49	**85**	**275**	67	**0.37**	0.42	**0.80**	**0.76**	**0.68**	**0.09**	0.64	**0.18**
SOM	Test	C12	0.69	**51**	102	258	**65**	0.33	**0.44**	**0.80**	0.72	0.65	0.08	**0.68**	0.14
Logit	Test	C18	0.72	77	**57**	**249**	93	**0.57**	**0.45**	0.73	**0.81**	0.68	0.13	0.72	0.28
SOM	Test	C18	0.60	**112**	89	217	**58**	0.56	**0.66**	**0.79**	0.71	**0.69**	**0.18**	**0.75**	**0.36**
Logit	Test	C24	0.58	132	68	185	91	0.66	0.59	0.67	0.73	0.67	0.16	0.76	0.33
SOM	Test	C24	0.63	**150**	**51**	**202**	**73**	**0.75**	**0.67**	**0.73**	**0.80**	**0.74**	**0.24**	**0.80**	**0.48**

Notes: See notes for Table 5.

Figure 1: The two-dimensional grid of the SOFSM

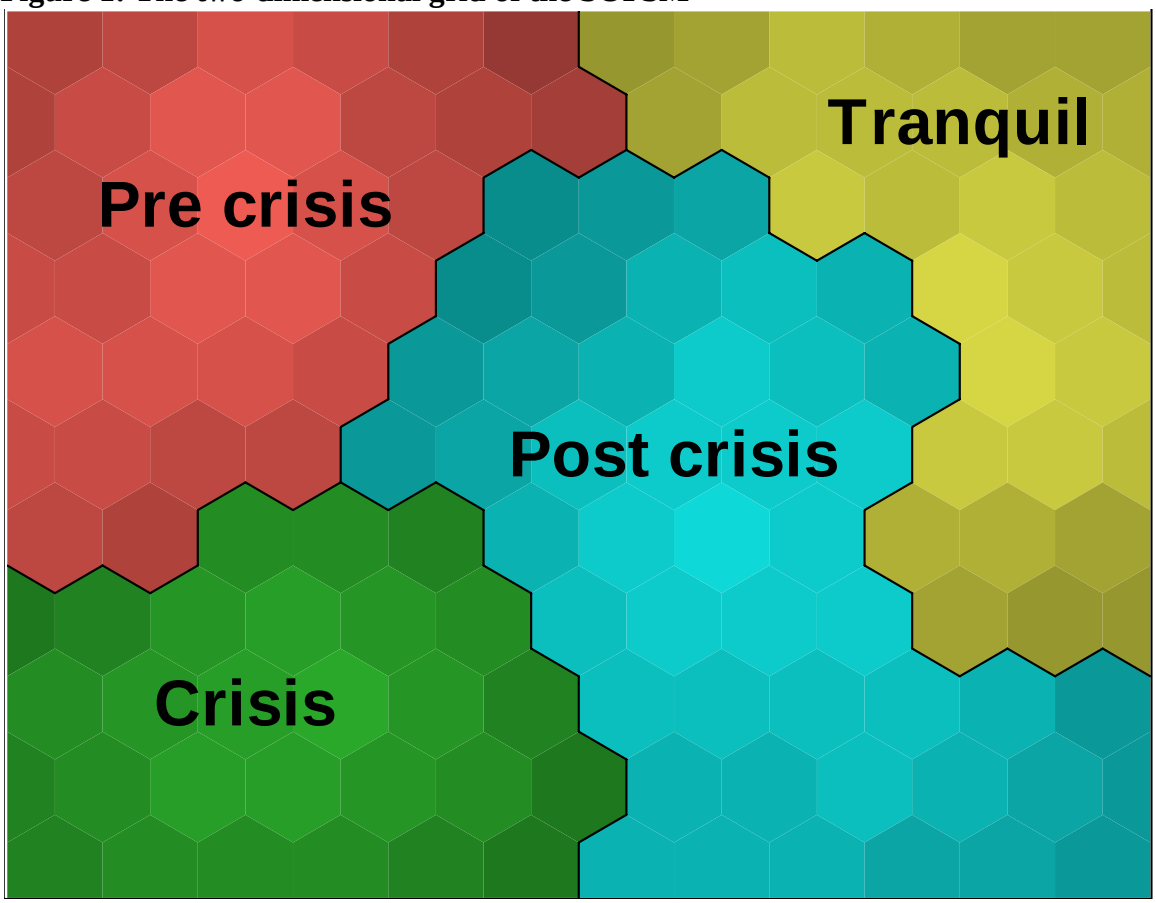
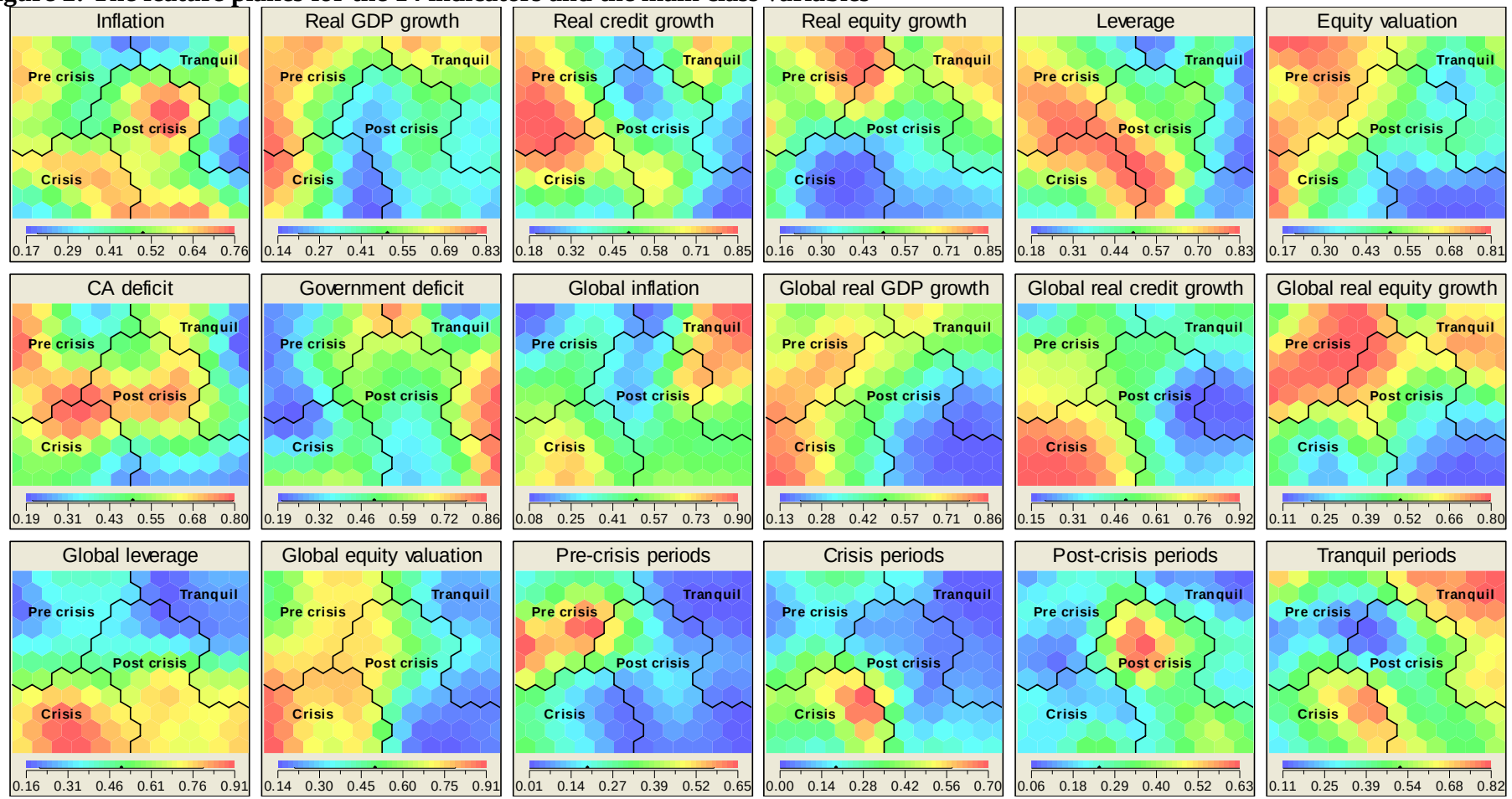
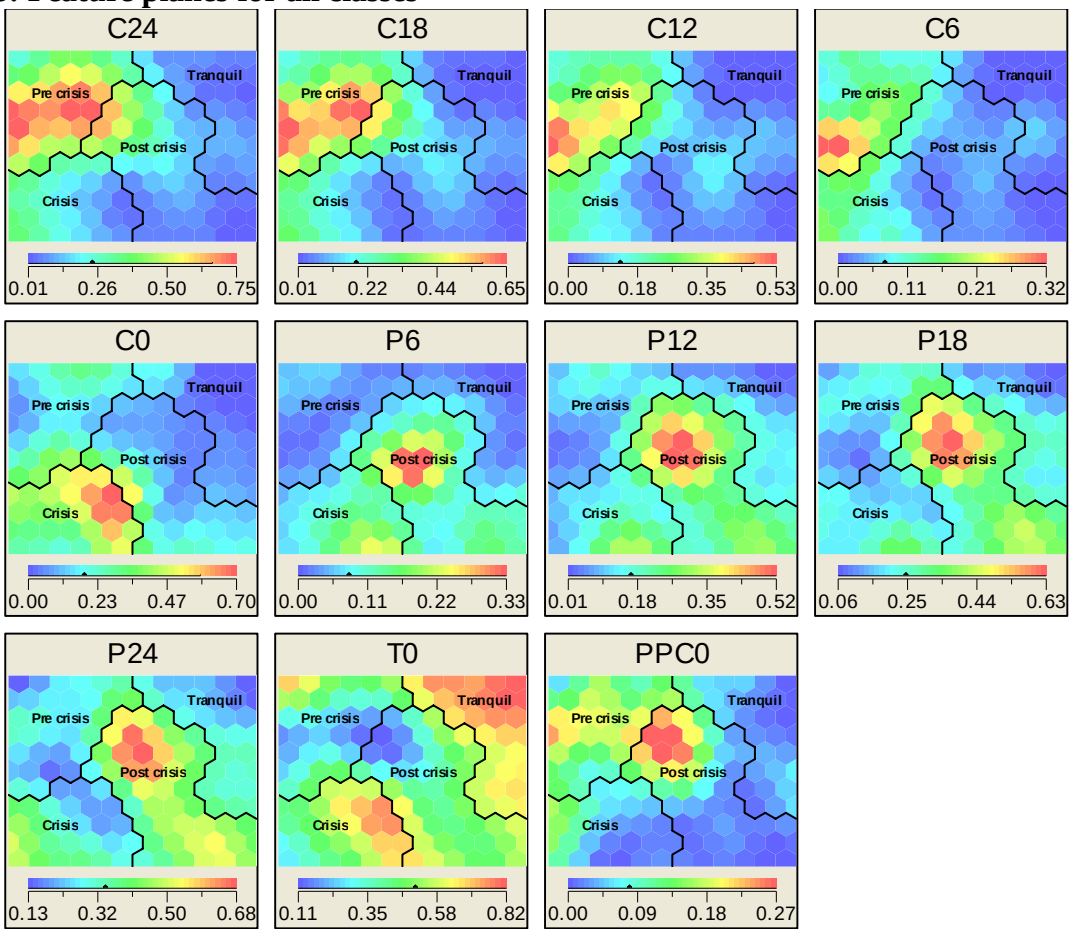


Figure 2: The feature planes for the 14 indicators and the main class variables



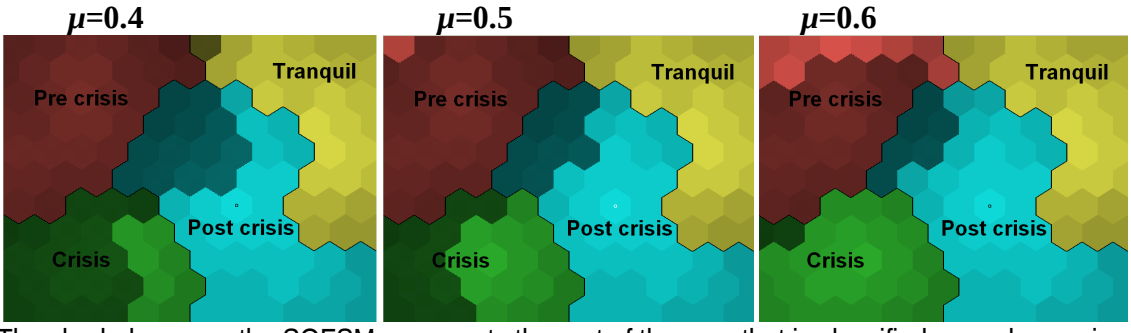
Notes: The feature planes are layers of the SOFSM in Figure 1 and show the distribution of each indicator on the grid. While the indicators are defined in Table 1, the four class variables are C18, C0, P18 and T0.

Figure 3: Feature planes for all classes



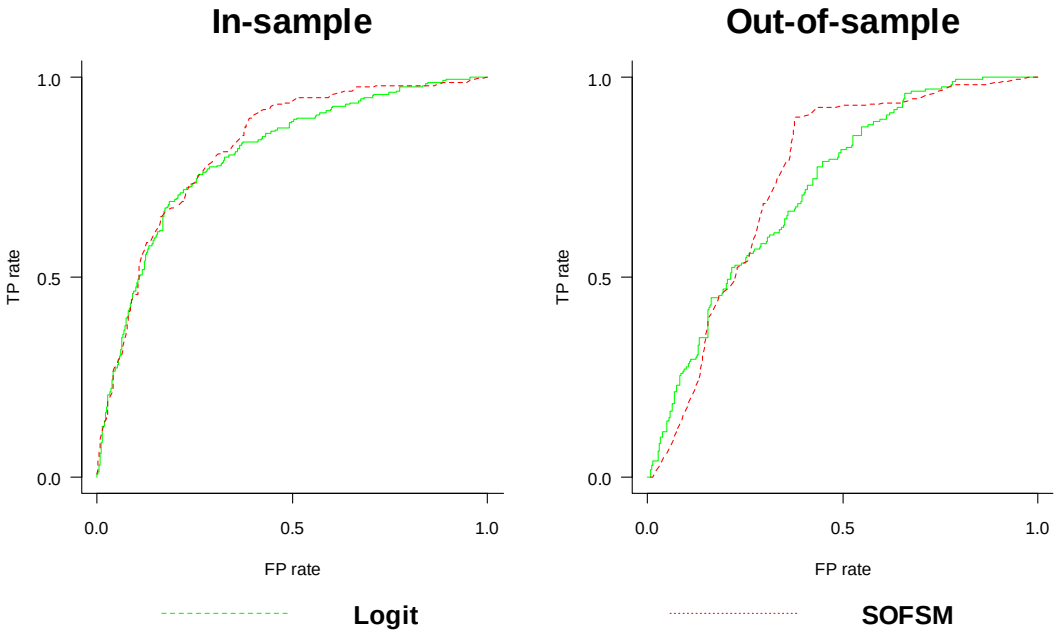
Notes: The feature planes C24, C18, C12, C6, P24, P18, P12 and P6 show the map distribution of class variables that represent 24, 18, 12 and 6 months before and after a crisis. C0 and T0 show the distribution of crisis and tranquil periods on the map.

Figure 4: Early warning nodes for different policymakers' preferences



Notes: The shaded area on the SOFSM represents the part of the map that is classified as early warning nodes when maximizing the policymakers' preferences with a horizon of 6 quarters according to the evaluation framework.

Figure 5: ROC curves for both models (μ=0.5 and horizon 6 quarters)



Notes: The vertical and horizontal axes represent TP rate (TP / (TP + FN)) and FP rate (FP/(FP+TN)). The AUC measures the area below these curves.

Figure 6: The evolution of the US and the Euro area

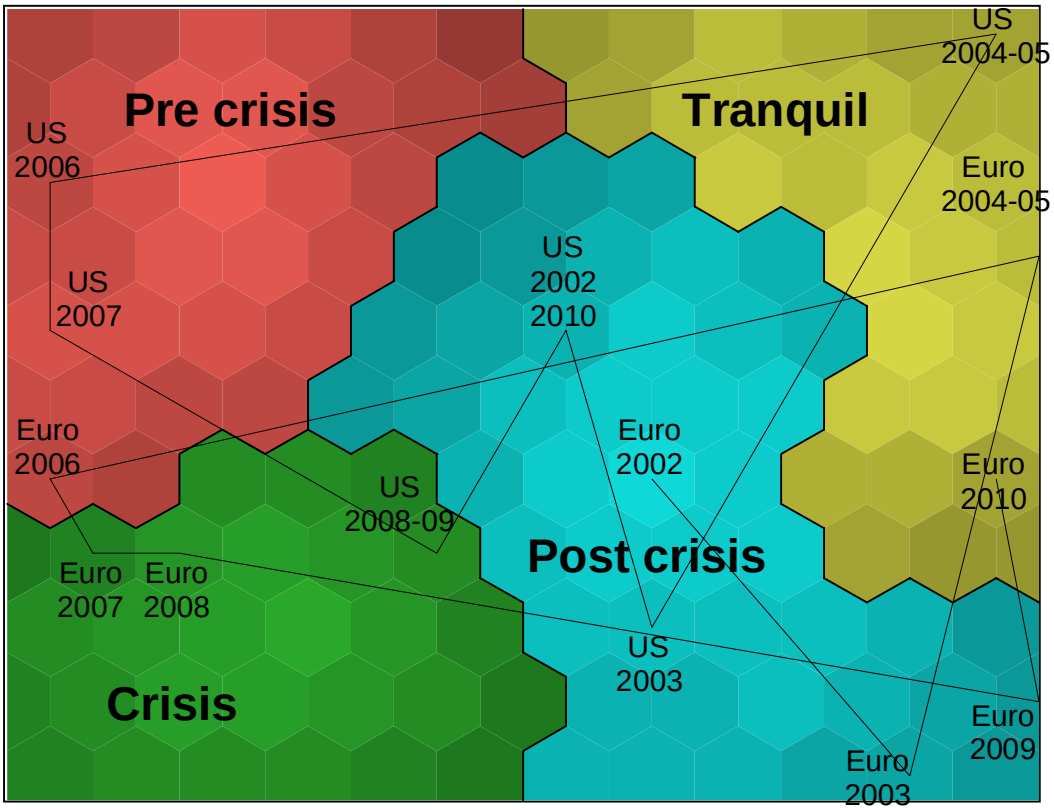


Figure 7: A mapping of all countries in 2010Q3

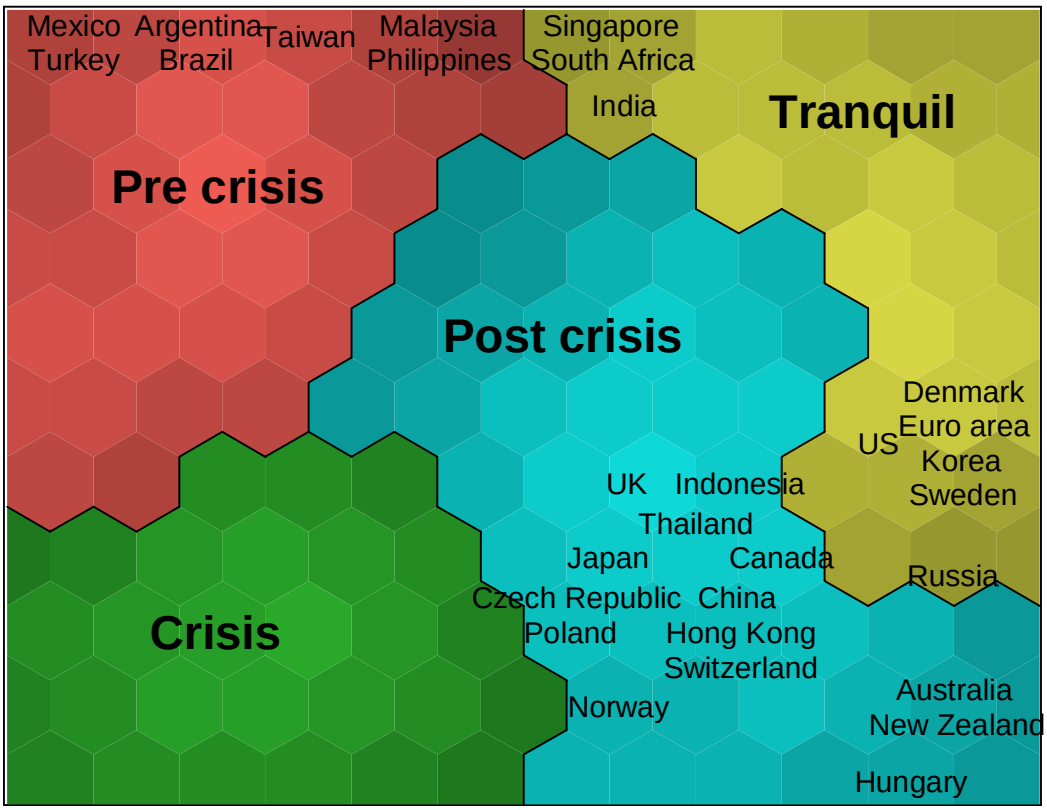


Figure 8: A mapping of the world aggregate from 2002–2010

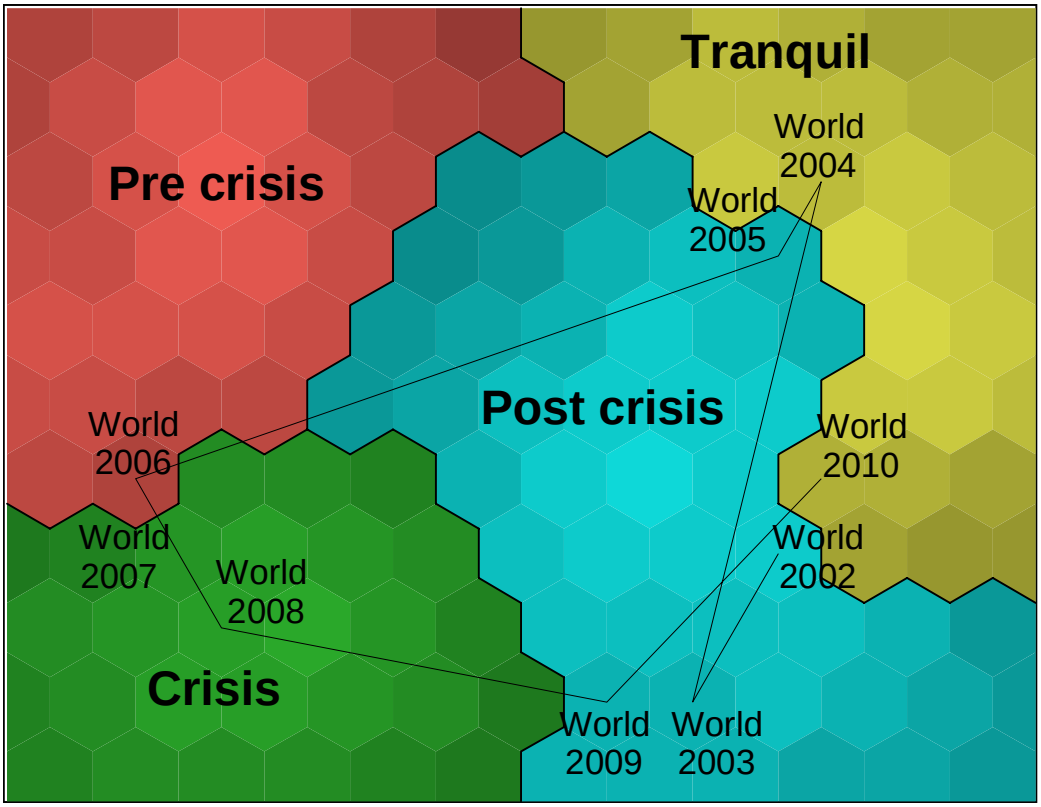
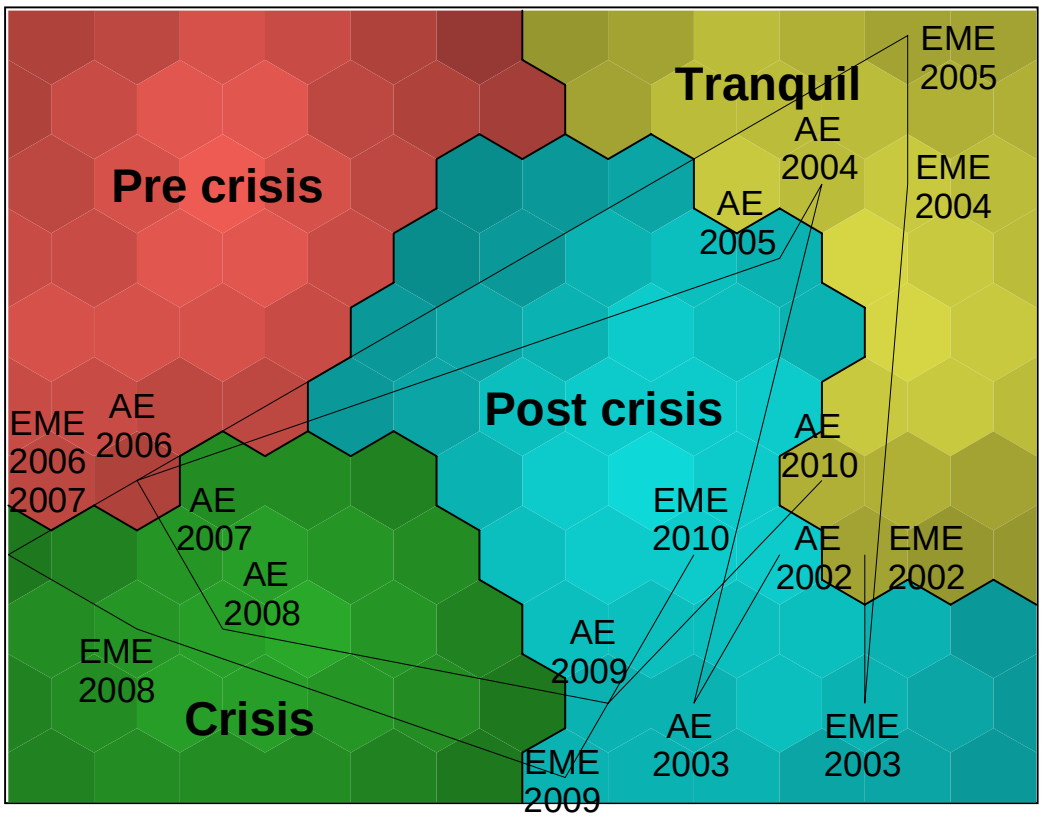


Figure 9: A mapping of advanced and emerging market economies' aggregates from 2002–10



Earlier BOFIT Discussion Papers

For a complete list of Discussion Papers published by BOFIT, see www.bof.fi/bofit

2010
- No 1 Anatoly Peresetsky: Bank cost efficiency in Kazakhstan and Russia
- No 2 Laurent Weill: Do Islamic banks have greater market power?
- No 3 Zuzana Fungáčová, Laura Solanko and Laurent Weill: Market power in the Russian banking industry
- No 4 Allen N. Berger, Iftekhar Hasan and Mingming Zhou: The effects of focus versus diversification on bank performance: Evidence from Chinese banks
- No 5 William Pyle and Laura Solanko: The composition and interests of Russia's business lobbies: A test of Olson's "encompassing organization" hypothesis
- No 6 Yu-Fu Chen, Michael Funke and Nicole Glanemann: Off-the-record target zones: Theory with an application to Hong Kong's currency board
- No 7 Vladimir Sokolov: Bi-currency versus single-currency targeting: Lessons from the Russian experience
- No 8 Alexei Karas, William Pyle and Koen Schoors: The effect of deposit insurance on market discipline: Evidence from a natural experiment on deposit flows
- No 9 Allen N. Berger, Iftekhar Hasan, Iikka Korhonen, Mingming Zhou: Does diversification increase or decrease bank risk and performance? Evidence on diversification and the risk-return tradeoff in banking
- No 10 Aaron Mehrotra and José R. Sánchez-Fung: China's monetary policy and the exchange rate
- No 11 Michael Funke and Hao Yu: The emergence and spatial distribution of Chinese seaport cities
- No 12 Alexey A. Ponomarenko and Sergey A. Vlasov: Russian fiscal policy during the financial crisis
- No 13 Aaron Mehrotra and Alexey A. Ponomarenko: Wealth effects and Russian money demand
- No 14 Asel Isakova: Currency substitution in the economies of Central Asia: How much does it cost?
- No 15 Eric Girardin and Konstantin A. Kholodilin: How helpful are spatial effects in forecasting the growth of Chinese provinces?
- No 16 Christophe J. Godlewski, Zuzana Fungáčová and Laurent Weill: Stock market reaction to debt financing arrangements in Russia
- No 17 Zuzana Fungáčová, Laurent Weill, Mingming Zhou: Bank capital, liquidity creation and deposit insurance
- No 18 Tuuli Koivu: Monetary policy, asset prices and consumption in China
- No 19 Michael Funke and Michael Paetz: What can an open-economy DSGE model tell us about Hong Kong's housing market?
- No 20 Pierre Pessarossi, Christophe J. Godlewski and Laurent Weill: Foreign bank lending and information asymmetries in China

2011
- No 1 Aaron Mehrotra and Jenni Pääkkönen: Comparing China's GDP statistics with coincident indicators
- No 2 Marco Lo Duca and Tuomas Peltonen: Macro-financial vulnerabilities and future financial stress - Assessing systemic risks and predicting systemic events
- No 3 Sabine Herrmann and Dubravko Mihaljek: The determinants of cross-border bank flows to emerging markets: New empirical evidence on the spread of financial crises
- No 4 Rajeev K. Goel and Aaron Mehrotra: Financial settlement modes and corruption: Evidence from developed nations
- No 5 Aaron Mehrotra, Riikka Nuutilainen and Jenni Pääkkönen: Changing economic structures and impacts of shocks - evidence from a DSGE model for China
- No 6 Christophe J. Godlewski, Rima Turk-Ariss and Laurent Weill Do markets perceive sukuk and conventional bonds as different financing instruments?
- No 7 Petr Jakubik: Households' response to economic crisis
- No 8 Wing Thye Woo: China's economic growth engine: The likely types of hardware failure, software failure and power supply failure
- No 9 Juan Carlos and Carmen Broto: Flexible inflation targets, forex interventions and exchange rate volatility in emerging countries
- No 10 Andrei Yakovlev: State-business relations in Russia in the 2000s: From the capture model to a variety of exchange models?
- No 11 Olena Havrylchyk: The effect of foreign bank presence on firm entry and exit in transition economies
- No 12 Jesús Crespo Cuaresma, Markus Eller and Aaron Mehrotra: The Economic transmission of fiscal policy shocks from Western to Eastern Europe
- No 13 Yu-Fu Chen, Michael Funke and Aaron Mehrotra: What drives urban consumption in mainland China? The role of property price dynamics
- No 14 Mikael Mattlin and Matti Nojonen: Conditionality in Chinese bilateral lending
- No 15 John Knight and Wei Wang: China's macroeconomic imbalances: Causes and consequences
- No 16 Gregory C Chow, Changjiang Liu and Linlin Niu: Co-movements of Shanghai and New York stock prices by time-varying regressions
- No 17 Yanbing Zhang, Xiuping Hua, Liang Zhao: Monetary policy and housing prices; a case study of Chinese experience in 1999-2010
- No 18 Peter Sarlin, Tuomas A. Peltonen: Mapping the state of financial stability

Bank of Finland
BOFIT – Institute for Economies in Transition
PO Box 160
FIN-00101 Helsinki

☎ + 358 10 831 2268
bofit@bof.fi
http://www.bof.fi/bofit